

ISSN 1936-5330

THE FEDERAL RESERVE BANK of KANSAS CITY
ECONOMIC RESEARCH DEPARTMENT

The Aggregate Implications of Individual Labor Supply Heterogeneity

José Mustre-del-Río

December 2011

RWP 11-09



RESEARCH WORKING PAPERS

The Aggregate Implications of Individual Labor Supply Heterogeneity *

José Mustre-del-Río[†]

Federal Reserve Bank of Kansas City

December 1, 2011

Abstract

This paper examines the Frisch elasticity at the extensive margin of labor supply in an economy consistent with the observed dispersion in average employment rates across individuals. An incomplete markets economy with indivisible labor is presented where agents differ in their disutility of labor and market skills. The model's key parameters are estimated using indirect inference with panel data from the National Longitudinal Survey of the Youth-NLSY. The estimated model implies an elasticity of aggregate employment of 0.71. A simple decomposition reveals that labor disutility differences, which capture the dispersion in employment rates, are crucial for this quantitative result. These differences alone generate an elasticity of 0.69. Meanwhile, skill differences alone imply an elasticity of 1.1. These results suggest that the literature generates large employment elasticities by ignoring individual labor supply differences.

*I would like to thank my advisor Mark Bills for his endless advice, comments and support on this project. I would also like to thank Mark Aguiar, William Hawkins and Ronni Pavan for their useful comments and discussions along with seminar participants at Rochester, the Federal Reserve Bank of Kansas City, Census Bureau, Congressional Budget Office, Auburn University, Kent State University and Carlos III. Finally, I am indebted to Yongsung Chang for his suggestions and help. The views expressed in this paper are solely mine and do not necessarily represent the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

[†]Email: jose.mustre-del-rio@kc.frb.org.

1 Introduction

The labor supply elasticity plays a crucial role in understanding employment fluctuations over the business cycle and in evaluating the effect of taxes and government spending. Early business cycle models, (e.g., Lucas and Rapping, 1969), require a representative agent to have a large intertemporal substitution of leisure to be consistent with the observed movements in hours and wages. Similarly, Prescott (2004) postulates a large aggregate elasticity of labor supply when determining the effect of marginal labor tax rates on labor supply across countries and time. Meanwhile, estimates based on labor supply decisions over the life-cycle find elasticities that are positive but economically small.¹

More recently, work by Chang and Kim (2007), Rogerson and Wallenius (2009), Gourio and Noual (2009), and Erosa et al. (2010) argues that one can generate a large macro elasticity in spite of assuming a small elasticity at the micro level. In these papers, the large employment response to wage changes is determined by differences across workers in the surplus that employment generates relative to nonemployment. Equivalently, the labor supply elasticity depends on how different are individual reservation wages compared to the market wage.

This paper measures the Frisch elasticity at the extensive margin of labor supply when individuals are ex-ante different in labor supply and skills, and hence heterogeneous in the surplus that employment generates for them. This heterogeneity is motivated by observations from data on individuals (National Longitudinal Survey of the Youth–NLSY) that show large differences in average employment rates that do not project on wages. A model is presented that is consistent with these facts. The model is a heterogeneous agent economy with incomplete markets and indivisible labor supply with two novel features. First, agents differ in their disutility of labor and second, they differ in their market

¹See for example Ghez and Becker (1975), MaCurdy (1981), Altonji (1986), or Abowd and Card (1989).

skills.

To impose quantitative discipline on the model, its key parameters are estimated with data from the NLSY using indirect inference. Due to the NLSY's structure long individual employment and wage histories can be constructed making it ideal for the present analysis. Moreover, because of its retrospective nature, data can be constructed at a quarterly frequency. This circumvents the time-aggregation issues that arise when using annual surveys, such as the Panel Study of Income Dynamics (PSID), as discussed in Erosa et al. (2010).

The result of the paper is that once agents display a realistic amount of ex-ante heterogeneity in labor supply, as well as wages, a very large macro-level elasticity is no longer obtained through the extensive margin of labor supply. The implied aggregate labor supply elasticity of the baseline model is 0.71. Robustness exercises generate elasticities as low as 0.62 and suggest that as the degree of labor supply heterogeneity in the economy increases, the Frisch elasticity at the extensive margin decreases. This value is below that reported in the literature, which typically generates extensive margin elasticities above 1.² At the same time, this elasticity is above estimates of the Frisch elasticity of the intensive margin of labor supply.³

Further inspection of the model reveals that labor disutility differences across agents are essential in generating the low labor supply elasticity. In a version of the baseline model with only ex-ante skill differences (in the spirit of Erosa et al., 2010), the implied elasticity is 1.1. Meanwhile, in a version of the model with only ex-ante labor disutility differences, the implied elasticity drops to 0.69.

²See for example Chang and Kim (2006, 2007), and Gourio and Noul (2009).

³This is consistent with the fact that over the business cycle the majority of employment adjustment occurs through extensive margin adjustments (see for example Heckman 1984 and Coleman 1984). Chetty (2010) finds estimates for the Hicksian elasticity of the intensive margin ranging from 0.47 to 0.54. He argues that for plausible parameter values the Frisch elasticity has a similar range. Chetty et al. (2011) find a lower bound for the elasticity at the intensive margin of 0.34. Meanwhile, Chetty et al. (2009) argue that the Frisch elasticity at the intensive margin is at most 0.63. Finally, Faberman (2010) finds intensive-margin elasticities ranging from 0.4 to essentially zero.

This version, however, generates a counterfactual wealth effect on participation. Similar to Chang and Kim (2007), in this version of the model the wealthiest do not participate in the labor market as much as in the data. Once labor disutility and skills are both incorporated, an elasticity of 0.71 is recovered. This complete model also generates a realistic wealth effect on participation.

This paper extends the literature that applies the neoclassical growth model to account for choices at the extensive margin of labor supply. In a representative agent model with indivisible labor, Hansen (1985) and Rogerson (1988) are the first to show that individual and aggregate labor supply elasticities are unrelated. Because of the representative agent assumption, there is no heterogeneity in markets skills or the value of nonmarket time. Hence, the Frisch elasticity at the extensive margin is infinite. Cho (1995) relaxes the representative agent assumption and allows for ex-post heterogeneity across agents in their market productivity. However, he maintains the complete markets assumption and assumes that market and nonmarket skills are correlated. This assumption preserves the infinite Frisch elasticity at the extensive margin.

Chang and Kim (2006) go a step further and relax both the representative agent and complete markets assumptions. Their model features ex-post heterogeneity as in Cho (1995), but no consumption insurance across agents. In their model, as in mine, the slope of the aggregate labor supply schedule is determined by the distribution of reservation wages. Their model does not allow for any ex-ante heterogeneity across agents and implies an elasticity at the extensive margin around 1.⁴ Finally, Gourio and Noual (2009) consider a model with complete markets where agents are ex-post heterogeneous in their labor productivity and taste for leisure. They estimate their model on data from the NLSY and obtain an aggregate elasticity of 1.3.

⁴Krusell et al. (2011) incorporate labor market frictions in a model similar to Chang and Kim (2006). Their model produces empirically reasonable patterns for transitions between employment and nonemployment if idiosyncratic shocks are persistent enough.

Rogerson and Wallenius (2009), and Erosa et al. (2010) adopt a different approach and examine life-cycle models. In these models a nonlinear mapping between hours of work and earnings plays a crucial role in providing the disconnect between micro and macro elasticities of labor supply. Both models allow for intensive and extensive margin adjustments. The work by Erosa et al. (2010) differs from Rogerson and Wallenius (2009) by allowing for incomplete markets and heterogeneous agents. In addition, Erosa et al. (2010) allow for ex-ante differences in skills across agents. Their model implies an elasticity at the extensive margin of 0.69, which is very close to the baseline aggregate (male and female) Frisch elasticity reported in this paper. Their theory abstracts from persistent differences in labor supply across individuals. This leads it to under-predict the variation in lifetime labor supply across individuals observed in the data, which is the focus of this paper. However, since they model life-cycle differences and allow for both intensive and extensive margin adjustments their model provides a wider subset of predictions consistent with individual level data.

While these contributions allow for heterogeneity across workers, the key dimension of ex-ante heterogeneity they lack is in the value of nonmarket time. This dimension of heterogeneity matters greatly for the Frisch elasticity at the extensive margin and is crucial to capture the average employment rate differences observed in the data. In the NLSY, most individuals are typically employed and therefore display high average employment rates. Meanwhile, others are employed less frequently and display relatively low average employment rates. This suggests fewer individuals are located at the margin than what is implied in the work of Chang and Kim (2007) or Gourio and Noual (2009).

In a model without ex-ante differences in labor supply, all individuals are, on average, employed at the same frequency and thus display similar employment rates. In equilibrium, since everyone's willingness to work is roughly the same, the reservation wage distribution is dense around the market wage. Thus, for a

small change in the wage rate there is a large aggregate labor supply response simply through individual extensive margin adjustments.

In the current model with ex-ante differences in labor supply, individuals differ in their average employment rates. Because of these labor supply differences, the reservation wage distribution implied by the estimated model is dispersed in a neighborhood around the equilibrium wage rate. As a result, for a small change in the wage rate there is a small aggregate labor supply response as few individuals change their employment decision due to the location of their reservation wage relative to the equilibrium wage.

This paper proceeds as follows. The next section describes the model. Section 3 presents the NLSY sample used for the empirical analysis. Section 4 discusses the estimation procedure. Section 5 presents the results of the estimation procedure along with a discussion of the model's fit to the data. Section 6 presents the implied Frisch elasticity of the estimated model and the decomposition of this elasticity. Section 7 presents robustness checks on the implied Frisch elasticity. Section 8 concludes.

2 Model

The model economy is a heterogeneous agent model with incomplete markets and indivisible labor supply similar to the one considered by Chang and Kim (2007). Unlike their work, agents are both ex-ante and ex-post heterogeneous. As in Erosa et al. (2010), individuals are ex-ante heterogeneous in skills. In addition, agents are also ex-ante heterogeneous in labor disutility, which is the key distinguishing feature of this model from the rest of the literature.⁵ These

⁵The two dimensions of heterogeneity across agents could alternatively be interpreted as market and non-market skills (as in Bils et al., 2009), leaving agents with a choice between working in the market or working at home. Because the data used to estimate the model's parameters has no information on non-market activities, it is not possible to distinguish between somebody valuing leisure more and working less in the market versus being more productive

two new dimensions of heterogeneity allow the model to account for differences across workers in average employment rates and wages. As in Aiyagari (1994), individuals are ex-post different in wealth and labor productivity. The analysis is confined to a steady-state with no aggregate uncertainty.

2.1 Workers

The economy is populated by a continuum (measure one) of workers. Workers differ in terms of their time invariant disutility of labor $d_j \in \{d_1, d_2, \dots, d_M\}$ and market skills $s_i \in \{s_1, s_2, \dots, s_N\}$. They also differ in their idiosyncratic productivity x that evolves exogenously according to the stochastic process with transition probability function $\pi_x(x'|x) = \Pr(x_{t+1} \leq x' | x_t = x)$. Workers have preferences over consumption c_t given by $\ln(c_t)$ to support a balanced growth path.

Workers can trade claims for physical capital a_t , which yields a rate of return r . Physical capital is the only asset available to workers (markets are incomplete) and they face a borrowing constraint $a_t \geq \bar{a}$ for all t as in Aiyagari (1994). Labor supply is indivisible as in Hansen (1985); Rogerson (1988). When employed, a worker with skills s_i must supply \bar{h} units of labor and earns $w_t x_t s_i \bar{h}$, where w_t is the market wage rate per unit of effective labor $x_t s_i$.

The value function of an employed worker with market skills s_i , disutility of labor d_j , assets a , and idiosyncratic productivity x is:

$$V_{ij}^E(a, x) = \max_{a'} \ln(c) - d_j + \beta \mathbf{E}[\max\{V_{ij}^E(a', x'), V_{ij}^{NE}(a', x')\} | x] \quad (1)$$

at home and working less in the market. Moreover, even if the data did include information on non-market work, the individual labor assumption in the model precludes the marginal decision between an hour of work in the market versus at home. Thus for expositional simplicity, the former interpretation is maintained.

subject to

$$\begin{aligned} c &= w s_i \bar{h} + (1 + r)a - a' \\ a' &\geq -\bar{a}. \end{aligned}$$

A worker takes the wage w and the interest rate r as given. Meanwhile, the value function of a non-employed worker is defined as:

(2)

$$V_{ij}^{NE}(a, x) = \max_{a'} \ln(c) + \beta \mathbf{E}[\max\{V_{ij}^E(a', x'), V_{ij}^{NE}(a', x')\} | x]$$

subject to

$$\begin{aligned} c &= (1 + r)a - a' \\ a' &\geq -\bar{a} \end{aligned}$$

Finally, the labor supply decision h of an individual with market skills s_i , disutility of labor d_j , assets a and idiosyncratic productivity x is thus characterized by:

$$V_{ij}(a, x) = \max_{h \in \{0, \bar{h}\}} \{V_{ij}^E(a, x), V_{ij}^{NE}(a, x)\}. \quad (3)$$

Note that the reservation productivity $x_{ij}^*(a)$, the value of x such that the worker is indifferent between working and not working, is an increasing function of asset holdings a and labor disutility d , but decreasing in market skills s . Because workers face the same stochastic process for x , differences in labor disutility will lead to systematic differences in the frequency of employment across workers, as low d workers will have a wider range of acceptable x 's and thus will be employed more often, relative to high d workers. Conditional on being em-

ployed at the same productivity level, high skill workers will also systematically earn higher wages relative to low skill workers, through the scaling effect of s_i on effective wages ws_i . It is thus through these two channels that the model will generate differences both in average employment and wages across workers. Meanwhile, the cross-sectional correlation between market skills and disutility of labor implicitly generates a cross-sectional correlation between average wages and employment.

The model abstracts from the intensive margin choice of labor supply and focuses on the extensive margin for several reasons. First, workers are rarely allowed to choose completely flexible work schedules or to supply a small number of hours. Second, a large fraction of hours fluctuations are accounted for by movements in and out of employment by workers (see for example Coleman, 1984; Heckman, 1984). Finally, Kimmel and Kniesner (1998) find that employment fluctuations account for three-fourths of wage-induced variation in labor hours.

Unlike Rogerson and Wallenius (2009), and Erosa et al. (2010), this model departs from the life-cycle. Instead, it follows the tradition of infinite horizon indivisible labor economies pioneered by Hansen (1985); Rogerson (1988) and continued by Chang and Kim (2006). As documented by Erosa et al. (2010), there is an inverted U-shape pattern in the participation of men over the life-cycle. Thus, this model potentially misses the elastic participation decisions of young and old individuals.⁶ This in principle, may lead the model to underestimate the true elasticity at the extensive margin. Note, however, that the elastic decision for both of these groups is fundamentally driven by a high value of non-market time (remaining in school or retiring). Hence, the model implicitly captures their age heterogeneity through high disutility of labor. Moreover, as shown in the next section, aggregate time-series of the extensive margin of

⁶Keane and Rogerson (2011) argue that the work of Keane and Wolpin (2000) implies an elasticity for young black males of 1.4. French (2005) finds a labor supply elasticity of 1.1 for men at age 60.

employment for prime-age workers (ages 25-54) and all workers display similar patterns.

2.2 Firms

There is a representative firm that takes capital K and effective units of labor L as inputs, and produces output Y according to a constant returns-to-scale Cobb-Douglas technology:

$$Y = F(K, L) = K^\alpha L^{1-\alpha} \quad (4)$$

Capital depreciates at a constant rate δ , while effective units of labor are measured as

$$L = \sum_{i=1}^N \sum_{j=1}^M \int h_{ij}(a, x) x s_i d\mu_{ij}. \quad (5)$$

Here $h_{ij}(a, x)$ is the labor supply decision of a worker of type $s = s_i, d = d_j$ with assets a and idiosyncratic productivity x ; $\mu_{ij} = \mu_{ij}(a, x)$ is the distribution of these workers. It is such that $\int d\mu_{ij} = p_{ij}$ and $\sum_{ij} p_{ij} = 1$, where p_{ij} denotes the proportion of workers with skills s_i and disutility of labor d_j .

2.3 Equilibrium

A steady-state equilibrium consists of a set of value functions $\{V_{ij}^E(a, x), V_{ij}^{NE}(a, x), V_{ij}(a, x)\}_{i=1, j=1}^{N, M}$, decision rules for consumption, asset holdings and labor supply, $\{c_{ij}(a, x), a'_{ij}(a, x), h_{ij}(a, x)\}_{i=1, j=1}^{N, M}$; aggregate inputs, K, L and factor prices w, r such that:

1. Individuals optimize: Given prices w and r , the individual decision rules $\{c_{ij}(a, x), a'_{ij}(a, x), h_{ij}(a, x)\}_{i=1, j=1}^{N, M}$ solve $\{V_{ij}^E(a, x), V_{ij}^{NE}(a, x), V_{ij}(a, x)\}_{i=1, j=1}^{N, M}$.

2. The representative firm maximizes profits:

- $w = F_2(K, L)$
- $r = F_1(K, L) - \delta$

3. The good market clears:

$$\sum_{i=1}^N \sum_{j=1}^M \int \{a'_{ij}(a, x) + c_{ij}(a, x)\} d\mu_{ij} = F(K, L) + (1 - \delta)K$$

4. Factor markets clear:

$$\begin{aligned} L &= \sum_{i=1}^N \sum_{j=1}^M \int h_{ij}(a, x) x s_i d\mu_{ij} \\ K &= \sum_{i=1}^N \sum_{j=1}^M \int a d\mu_{ij} \end{aligned}$$

5. Individual and aggregate behaviors are consistent: For all $A^0 \subset \mathcal{A}$ and $X^0 \subset \mathcal{X}$ and each i, j ,⁷

$$\mu_{ij}(A^0, X^0) = \int_{A^0, X^0} \left\{ \int_{\mathcal{A}, \mathcal{X}} \mathbf{1}_{a'=a'_{ij}(a, x)} d\pi_x(x'|x) d\mu_{ij} \right\} da' dx'$$

3 Data

The data used comes from the National Longitudinal Survey of Youth 1979 (NLSY79), survey years 1990 through 2000. The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when first interviewed in 1979. Interviews were conducted annually through 1994 and biennially thereafter. Participants are asked questions regarding their family background, education, and work experience. Since average hourly wages

⁷Let \mathcal{A} and \mathcal{X} denote the sets of all possible realizations of a and x , respectively.

and employment rates are the primary focus of this study, the NLSY is used as it consistently tracks workers' employment histories over several years. While individuals are not interviewed on a quarterly basis, it is possible to convert the data to a quarterly frequency as individuals are asked information both on jobs currently held and held since the last interview including calendar dates on when each of the jobs started and finished.

Using quarterly information on employment and wages circumvents the bias introduced by time aggregation when using lower frequency data such as the PSID. This point is mentioned in Erosa et al. (2010), who argue that the wage rate obtained in the PSID as the ratio of annual earnings to annual hours is a noisy measure of the true returns to work faced by an individual during the year. This is because temporary low wage shocks will be unobserved in annual data if the individual chooses not to work during that portion of the year.

The drawback of using the NLSY is that respondents are fairly young when first interviewed in 1979. However, by the 1990 survey wave the youngest age reported by individuals is 26. Conversely, the oldest age observed in 2000 is 48. To gauge the representativeness of this age group for studying the responsiveness of labor force participation figure 1 compares the employment to population ratio for all workers and for workers ages 25-54. As can be seen from the figure this age group has an overall higher participation rate. This difference is driven mostly by the very low participation rate of workers near or in retirement. Projecting the HP filtered employment to population series of workers ages 25-54 on the HP filtered series of employment to population for all workers results in a coefficient of 1.04. This simple exercise suggests that workers in the 25-54 age group and workers from all age groups display very similar time-series movements at the extensive margin of employment.

The data is restricted to the cross-sectional subsamples in the NLSY. Individuals must not be in the armed forces and not be attending school. In addition,

over the 11 year period considered, individuals must have at least 22 quarters where employment status can be determined (either employed or non-employed).⁸ When employed, individuals must have data on both hours and wages earned. Jobs where hourly wages are below \$1.00 or above \$500 (in 1983 dollars) are ignored. Jobs where the individual works less than 30 hours a week are also ignored to restrict attention to full-time work. Finally, individuals must have non-zero average employment rates over the 11 year period considered. From the perspective of the model these individuals are not marginal, in the sense of being near the margin between choosing employment or non-employment. Hence, their exclusion should upwardly bias the implied elasticity as their elasticity is trivially 0. The resulting sample consists of 220,199 observations from 5,082 individuals. Summary statistics appear in table 1.

Table 1: Summary Statistics NLSY Panel 1990-2000.

Variable	Mean	Standard Deviation
Average employment rate	0.749	0.283
Average log wage	1.994	0.521
Employment duration (in quarters)	15.2	11.833
Pr(E→N)	0.031	0.173
Pr(N→E)	0.121	0.326
Age	34.26	3.8
Male	0.497	0.500
White	0.79	0.394
Highest Grade Completed	13.52	2.54

Notes: Wages are in 1983 dollars. Cross-sectional correlation between average employment rates and wages equals 0.394.

⁸For those individuals with missing observations, they must have at least 22 valid quarters before the first missing observation as valid observations after the first missing observation are ignored. This is done for simplicity as adding valid observations after the first missing observation only increases my sample size by 3%. Moreover, the model simulation becomes more complicated as simulated data must replicate the observed frequency of valid observations after the first missing observation.



Figure 1: Employment to Population Ratio, All Workers (solid) and Workers Ages 25-54 (dash) from CPS.

4 Model Parametrization and Estimation

4.1 Parametrization

This section describes how the model is parametrized and the procedure used to estimate its key structural parameters. Details of how the steady-state equilibrium is computed appear in the appendix. To start, the unit of time is a quarter. Individual productivity x follows an $AR(1)$ process: $\ln x' = \rho_x \ln x + \epsilon_x$, where $\epsilon_x \sim N(0, \sigma_x^2)$. As in Chang and Kim (2007), an employed individual spends one-third of discretionary time working, so $\bar{h} = \frac{1}{3}$. The capital-income share α is set to 0.36 while the depreciation rate δ is set to 2.5 percent. The discount factor β is chosen so that in equilibrium the quarterly rate of return on capital is 1 percent.

Market skills and labor disutility take on three values $\{s_1, s_2, s_3\}$ and $\{d_1, d_2, d_3\}$, yielding a total of 9 distinct worker types and hence 9 proportions p_{ij} to be de-

terminated.⁹ By normalization, the highest skill level s_3 is set to 1, while p_{13} is set so that $\sum_{ij} p_{ij} = 1$. Under these assumptions, there are a total of 15 structural parameters that must be estimated: $\Psi' = (s_1, \dots, d_1, \dots, p_{11}, \dots, \rho_x, \sigma_x)$. The procedure used to estimate these 15 parameters is discussed next.

4.2 Estimation via Indirect Inference

The vector of structural parameters Ψ is estimated using indirect inference rather than directly given the complicated structure of the model.¹⁰ Indirect inference involves the use of an auxiliary statistical model that serves as a criterion to determine if actual data and model-generated data (given Ψ) are “close enough” in a sense that is formally defined below. Define the indirect inference estimator of Ψ , as the estimated value $\hat{\Psi}$ that is found when the estimated parameters of the auxiliary model obtained when using actual data and the estimated parameters of the auxiliary model obtained when using model-simulated data are close enough.

More formally, suppose that the observed data can be written as $\{y_{it}\}, i = 1, \dots, N; t = 1, \dots, T$, while data generated from the model can be written as $\{\tilde{y}_{it}(\Psi)\}, i = 1, \dots, N; t = 1, \dots, T$.

Next, suppose the auxiliary model is characterized by a vector of parameters Γ (of dimension $p \geq k$) that can be estimated using observed data as:

$$\hat{\Gamma} = \arg \max_{\Gamma} \mathcal{L}(y; \Gamma), \quad (6)$$

where $\mathcal{L}(y; \Gamma)$, is the likelihood function associated with the auxiliary model.

Meanwhile, the model can be simulated to generate M statistically independent data sets $\{\tilde{y}_{it}^m(\Psi)\}, m = 1, \dots, M$. As in the case with observed data, the

⁹The choices for the number of skills and labor disutilities is primarily driven by computational concerns as adding more worker types increases the state-space and thus computational time significantly. As will be seen in the next section these modeling choices seem to fit the data well. However, section 7 relaxes this assumption.

¹⁰This method was first introduced by Smith (1990,1993) and extended by Gourieroux, Monfort, and Renault (1993) and Gallant and Tauchen (1996).

auxiliary model can be estimated using each of the simulated data sets to obtain M estimated parameter vectors $\tilde{\Gamma}_m(\Psi)$, as:

$$\tilde{\Gamma}_m(\Psi) = \arg \max_{\Gamma} \mathcal{L}(y^m(\Psi); \Gamma), \quad (7)$$

Finally, define the average of the estimated parameter vectors by $\tilde{\Gamma}(\Psi) = M^{-1} \sum_{m=1}^M \tilde{\Gamma}_m(\Psi)$. The criterion used to determine if the observed data and simulated data are “close enough” through the lens of the auxiliary model is the Wald approach to indirect inference that chooses Ψ to minimize the quadratic form in the vector $\hat{\Gamma} - \tilde{\Gamma}(\Psi)$:

$$\hat{\Psi}^{Wald} = \arg \min_{\Psi} (\hat{\Gamma} - \tilde{\Gamma}(\Psi))' W (\hat{\Gamma} - \tilde{\Gamma}(\Psi)) \quad (8)$$

where W is a positive definite “weighting” matrix.^{11,12}

Notice that accommodating sample restrictions and attrition when estimating Ψ via indirect inference is straight forward. One needs to apply the same sample restrictions and assumptions on attrition across actual and simulated data sets. In the present context, each simulated data set consists of $I = 5082$ individuals contributing at most 44 quarters of data, as in the panel constructed from the NLSY. Because some individuals have fewer quarterly observations due to attrition, simply omit quarter observations in the simulated data so that the distribution of “quarter-counts” by individual in model-generated data is the

¹¹For the purposes of this paper W is set to the identity matrix \mathbf{I}_p . More generally, the optimal weighting matrix is the inverse of the covariance matrix of the parameter vector $\hat{\Gamma}$ using observed data. Note that setting $W = \mathbf{I}_p$ only affects the efficiency of the estimated $\hat{\Psi}$, but not its consistency.

¹²In practice, a Nelder-Meade simplex algorithm is used to minimize (8), as implemented in Press et al. (1992), with $M = 20$. As highlighted by Smith (2008), the usage of simulations inflates asymptotic standard errors by a factor of $(1 + M^{-1})^{1/2}$, and thus for $M \geq 10$, this factor is negligible.

same as in actual data.

4.3 The Auxiliary model

The auxiliary model choice is driven by two considerations: efficiency and computational complexity. From the perspective of efficiency it is important that the auxiliary model be flexible enough to provide a good description of the data. As stressed by Keane and Smith (2003), if the auxiliary model is correctly specified (in the sense that it provides a correct statistical description of the observed data), then the Wald approach to indirect inference is asymptotically equivalent to maximum likelihood, provided that M is sufficiently large. From the perspective of computational complexity, the auxiliary model should be one that can be estimated quickly as its parameters must be estimated M times for each choice of the structural parameters Ψ . Guided by these two considerations and following the related literature (Keane and Smith, 2003; Altonji et al. 2009), the following system of seemingly unrelated regressions (SUR) is used:

$$\begin{aligned}
E_{it} \cdot E_{it-1} &= \gamma_0^{EE} + \gamma_{ED}^{EE} \ln(ED_{it-1} + 1) + \gamma_{ND}^{EE} \ln(ND_{it-1} + 1) + \gamma_w^{EE} w_{it-1}^* + \gamma_e^{EE} e_i + \gamma_w^{EE} w_i + \epsilon_{it}^{EE} \\
E_{it} \cdot (1 - E_{it-1}) &= \gamma_0^{EN} + \gamma_{ED}^{EN} \ln(ED_{it-1} + 1) + \gamma_{ND}^{EN} \ln(ND_{it-1} + 1) + \gamma_w^{EN} w_{it-1}^* + \gamma_e^{EN} e_i + \gamma_w^{EN} w_i + \epsilon_{it}^{EN} \\
(1 - E_{it}) \cdot E_{it-1} &= \gamma_0^{NE} + \gamma_{ED}^{NE} \ln(ED_{it-1} + 1) + \gamma_{ND}^{NE} \ln(ND_{it-1} + 1) + \gamma_w^{NE} w_{it-1}^* + \gamma_e^{NE} e_i + \gamma_w^{NE} w_i + \epsilon_{it}^{NE} \\
(1 - E_{it}) \cdot (1 - E_{it-1}) &= \gamma_0^{NN} + \gamma_{ED}^{NN} \ln(ED_{it-1} + 1) + \gamma_{ND}^{NN} \ln(ND_{it-1} + 1) + \gamma_w^{NN} w_{it-1}^* + \gamma_e^{NN} e_i + \gamma_w^{NN} w_i + \epsilon_{it}^{NN} \\
w_{it}^* &= \gamma_0^w + \gamma_{ED}^w \ln(ED_{it-1} + 1) + \gamma_{ND}^w \ln(ND_{it-1} + 1) + \gamma_w^w w_{it-1}^* + \gamma_e^w e_i + \gamma_w^w w_i + \epsilon_{it}^w
\end{aligned}$$

or more compactly:

$$Y_{it} = Z_{it}\Gamma + \epsilon_{it} \quad (9)$$

where $\epsilon_{it} \sim N(0, \Sigma)$ and iid over i and t . The variable E_{it} denotes individual i 's employment status (1 or 0) in period t ; ED_{it-1} denotes the number of periods individual i has been continuously employed up to time $t-1$; ND_{it-1} denotes the number of periods individual i has been continuously non-employed up to time $t-1$; e_i is the individual's average employment rate; and w_i is the individual's average log hourly wage rate (conditional on being employed).^{13,14} The variable w_{it}^* represents the individual's log wage that is equal to the sample mean when non-employed and equals the observed wage otherwise.¹⁵

The auxiliary model in (9) is intended to capture in a succinct fashion the joint dynamics of wages and employment. The first four equations represent the four possible employment transitions in the model (employment to employment, employment to non-employment, etc.), while the last equation models the evolution of individual wages. This system of equations is a variant of the auxiliary model used in Altonji et al. (2009).¹⁶ This is a natural starting point as the model they ultimately estimate using generalized indirect inference can be interpreted as a reduced-form version of the current structural model. Unlike the previous literature, this auxiliary model includes terms that explicitly capture permanent differences across agents as embodied by their average employment rates and

¹³Both ED_{it} and ND_{it} are determined recursively as $ED_{it} = E_{it}(ED_{it-1} + 1)$ and $ND_{it} = (1 - E_{it})(ND_{it-1} + 1)$, respectively.

¹⁴To control for age effects in the data, which are absent in the model, the auxiliary model is estimated separately for four age groups: [25, 30), [30, 35), [35, 40), [40, 48). Then, for each regression coefficient the age-corrected estimate is defined as the weighted average of this coefficient across age groups. This procedure is designed to capture the effect of each variable on the average individual. Experimentation reveals that the estimated coefficients do not vary by much across age groups and are robust to the number of age groups used.

¹⁵Setting the wage equal to the sample mean when non-employed is valid so long as this assumption is maintained both in the actual and simulated data. Alternatively, one could set the wage of non-employed workers to zero, as suggested in Keane and Smith (2003).

¹⁶The choice of restricting the covariates to be the same across equations is driven by computational simplicity as the SUR system can be estimated via equation-by-equation OLS.

wages. By this dimension, the closest work is Guvenen and Smith (2010) who use average income as an explanatory variable in their auxiliary model that is then used to estimate a consumption-savings model.

Note that the system described in (9) consists of 45 parameters: 30 coefficients from the five equations and 15 unique elements in the covariance matrix Σ . Given that the identification of the two dimensions of heterogeneity (labor disutility and skills) precisely comes from cross-sectional variation in average employment and wages, it seems valuable for the purposes of calculating the Frisch elasticity at the extensive margin to discipline the estimation of model parameters by having model-generated data imitate these two distributions and their correlation. To this end, the means' (μ_e, μ_w) , standard deviations' (σ_e, σ_w) , skewness' $(Skew_e, Skew_w)$, kurtosis' $(kurt_e, kurt_w)$, and cross-sectional correlation ρ_{ew} of the distributions of average employment rates and average wages are estimated from actual and model generated data. These additional parameters yield a total of 53 auxiliary parameters that are used to *indirectly infer* the 15 elements of Ψ .

5 Estimation Results

This section presents the estimation results. The estimated parameters of the model are presented first. Next, the goodness of fit of the model is discussed by showing how well it performs in replicating the motivating cross-sectional facts regarding average employment rates and wages.

5.1 Estimated Model Parameters

Table 2 presents the estimated values for Ψ , the vector of structural parameters of the model. Given that the highest skill level is normalized to 1, these estimates imply that in the model the lowest skill type is 67 percent less productive in the

market relative to the highest skill. The results imply much larger variation in terms of labor disutility. Given log preferences over consumption, a d_1 type worker requires a 23 percent increase in consumption to offset her disutility of labor. Likewise, d_2 and d_3 type workers require increases in consumption of 75 and 233 percent, respectively, to be indifferent between working and not. Thus, for type d_3 individuals work is significantly more costly, in consumption terms, relative to type d_1 workers.

The estimated persistence of the productivity process (0.92) is below the estimates for males and females that Chang and Kim (2006) estimate (0.948 and 0.925, respectively) using data from the PSID and below what Chang and Kim (2007) report for all workers (0.929). The estimated standard deviation of the innovations to the productivity shock is also lower in comparison (0.18 versus 0.269 for males, 0.319 for females, and 0.227 overall). Krusell et al. (2011) study a frictional search model similar to Chang and Kim (2007) and calibrate it to match the persistence of employment and out of the labor force states across individuals. They obtain a persistence parameter of 0.9931 and a standard deviation of 0.1017. Compared to Chang and Kim (2006) and Krusell et al. (2011), this model requires less persistence in idiosyncratic shocks because of the inclusion of permanent ex-ante differences in labor supply and skills. Compared to Chang and Kim (2006), controlling for these differences also helps explain wage variation and hence the estimated standard deviation of the idiosyncratic shocks falls. Relative to Krusell et al. (2011), the estimated standard deviation is higher as the current model abstract from search frictions, which help explain some of the observed wage variation across workers. Finally, the fact that most individuals are located along the diagonal of the matrix of disutility versus skill is expected given that the model must reproduce a positive correlation between employment (labor supply) and wages (skills).

Given the estimated parameter values from table 2, the aggregate steady-state

employment rate of the model is 76.1 percent. Table 3 presents the steady-state employment rates conditional on worker type. As expected given the utility specification, the model predicts fairly large employment rate differences across disutility types, and small differences within types. While the lowest disutility types are employed nearly all the time, the highest disutility types are employed roughly one quarter of the time. Hence, as in the data, most of the individuals in the model display very high average employment rates, while a few others display comparatively low average employment rates.

Table 2: Estimated Parameter Values

Disutility of labor	Skills		
	$s_3 = 1.00^\dagger$	$s_2 = 0.48$ (0.03)	$s_1 = 0.33$ (0.03)
$d_1 = 0.23$ (0.02)	$p_{13} = 0.34$ (0.02)	$p_{12} = 0.05$ (0.02)	$p_{11} = 0.01$ (0.02)
$d_2 = 0.75$ (0.02)	$p_{23} = 0.01$ (0.03)	$p_{22} = 0.42$ (0.03)	$p_{21} = 0.01$ (0.02)
$d_3 = 2.34$ (0.03)	$p_{33} = 0.00$ (0.02)	$p_{32} = 0.00$ (0.02)	$p_{31} = 0.16^\dagger$
Value			
ρ_x	0.92 (0.02)		
σ_ϵ	0.18 (0.02)		

Notes: † by normalization. Asymptotic standard errors appear in parentheses.
Discount factor $\beta = 0.98812$, found from capital market clearing.

Table 3: Model Steady-state Employment Rates, by Worker Type

Disutility of labor	Skills		
	s_3	s_2	s_1
d_1	0.99	0.99	0.99
d_2	0.72	0.74	0.75
d_3	0.23	0.23	0.25

Notes: Aggregate employment rate is 0.748.

5.2 Assessing the Model's Fit

This subsection discusses the goodness of fit of the model. First, the employment and wage distributions across model and actual data are presented. Second, this section shows how well the model replicates the negative duration dependence in the hazard rates out of and into employment observed in the data. Finally, a comparison of wealth distributions for both sources of data is discussed.

5.2.1 Employment and Wages

Figure 2 presents the distributions of average employment rates obtained from actual and model-generated data, while figure 3 presents the analogous distributions of average wages. Most striking from figure 2 is how well the model matches the data distribution of employment rates. However, the model over-predicts the portion of individuals with average employment rates near 100%. In terms of the distribution of wages, the model also performs well. Relative to the data the distribution of wages in the model is slightly less disperse as few individuals in the model earn very high wages.

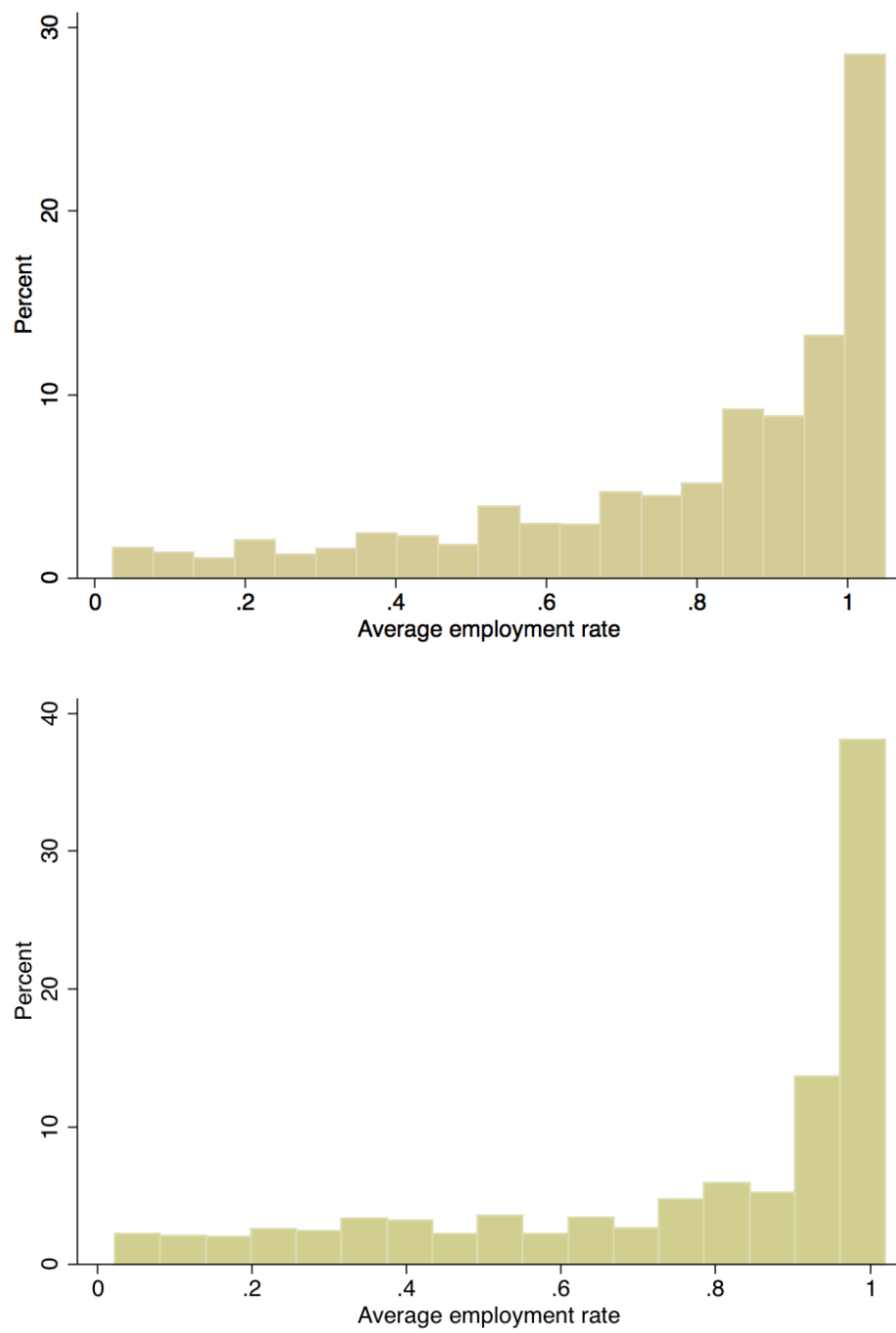


Figure 2: Distribution of Average Employment Rates, Data (top) and Model (bottom).

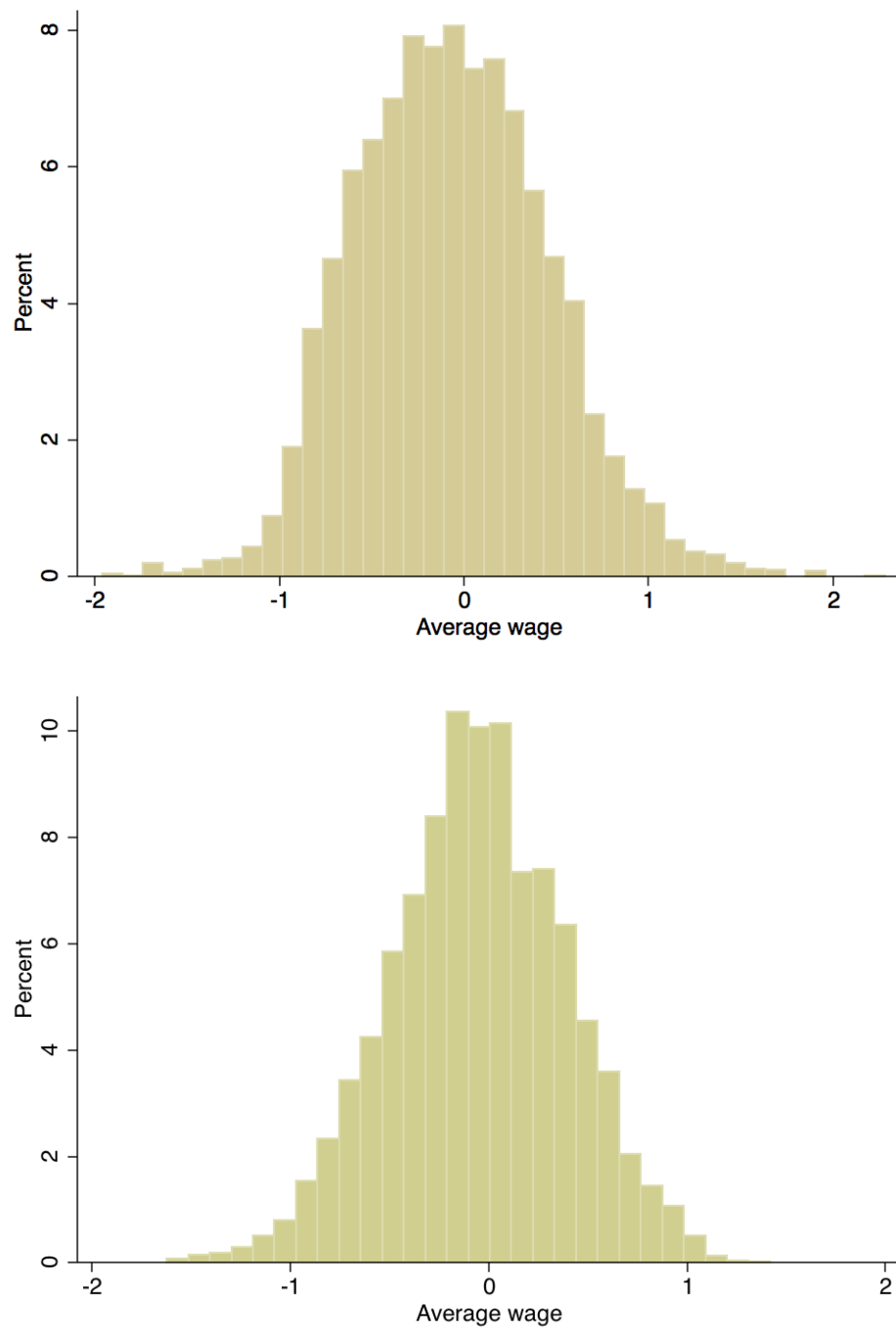


Figure 3: Distribution of Average Wage Rates, Data (top) and Model (bottom).

5.2.2 Hazard Rates

As can be seen from figures 4 and 5, the model is able to capture the negative duration dependence of both the hazard from employment to non-employment and the hazard from non-employment to employment. However, the model over-predicts the decline in both of these hazards for spells lasting at most 2 quarters. The reason for this result is purely compositional. In the model, flows from employment to non-employment occurring within the first 2 quarters of the duration of an employment spell, are disproportionately done by workers with the highest disutility of labor d_3 . Because these workers dislike market work so much, they engage in short lived employment spells, consistent with their low average employment rates. Likewise, flows from non-employment to employment occurring within the first 2 quarters of the duration of a non-employment spell also are disproportionately done by these same workers. Looking at the model's predicted hazards after 2 quarters (once most of the effect of type d_3 workers vanishes), the model performs better in replicating both the direction and level of both hazard rates.

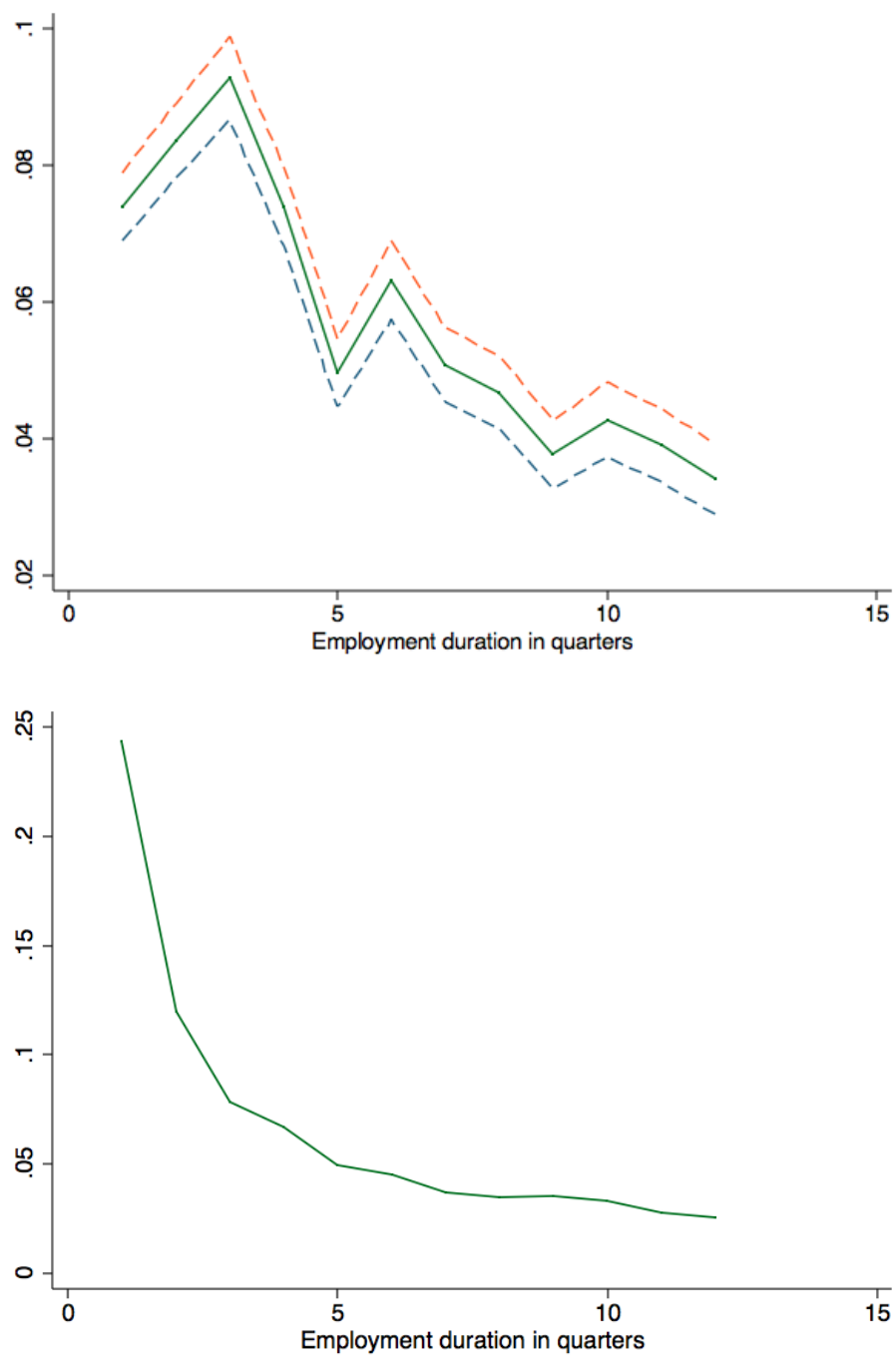


Figure 4: Hazard Rates from Employment to Non-employment, Data (top) and Model (bottom).

Note: Dashed lines represent 95% confidence interval of data.

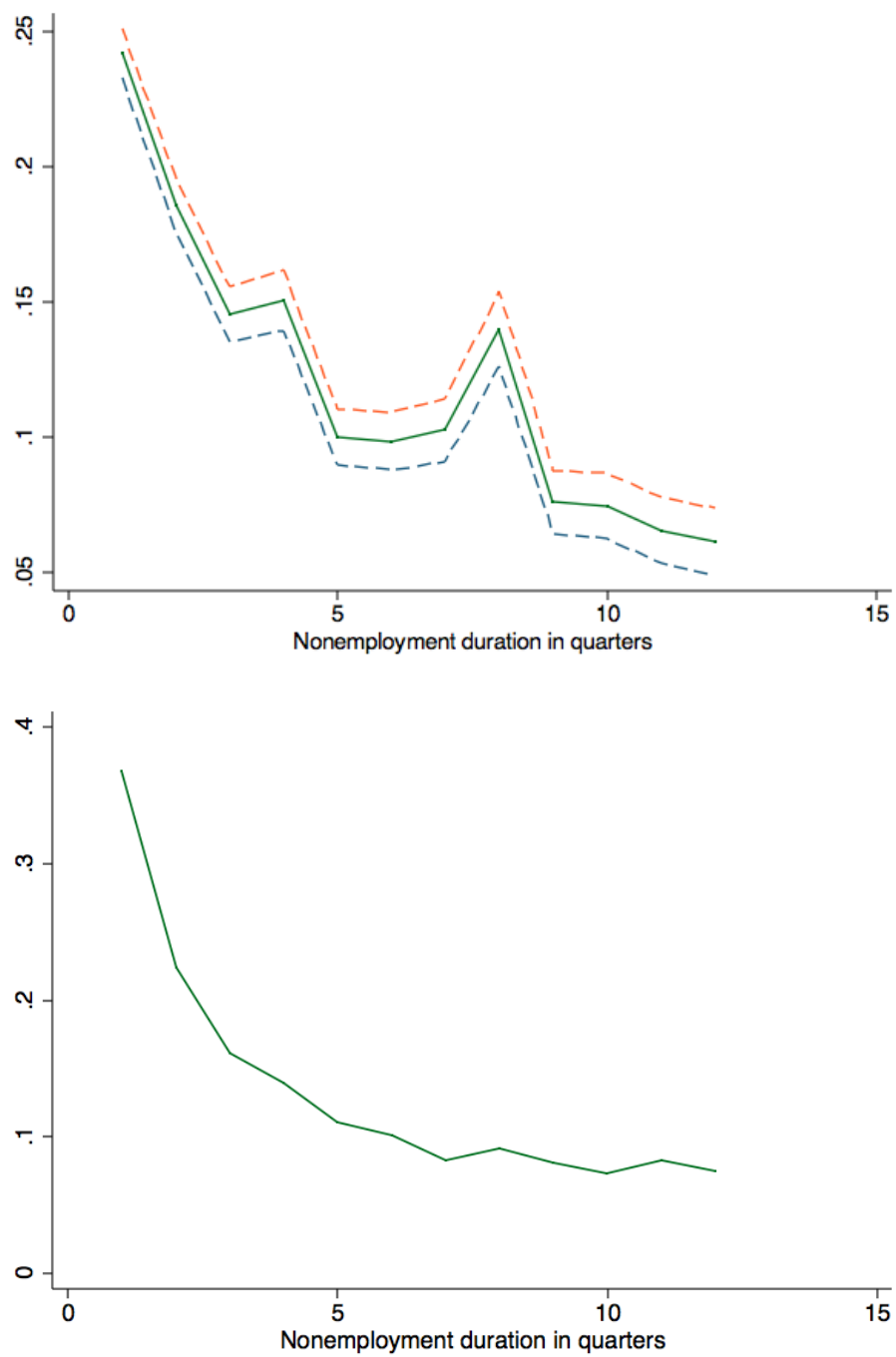


Figure 5: Hazard Rates from Non-employment to Employment, Data (top) and model (bottom).

Note: Dashed lines represent 95% confidence interval of data.

5.2.3 Wealth

As a final check on the model, this section examines how well it replicates the cross-sectional wealth and earnings distribution observed in the data. Table 4 presents detailed statistics on wealth and earnings from the PSID, the baseline model, and for comparison, Chang and Kim’s (2007) model. As in Chang and Kim (2007), the category “PSID Primary Households” reflects households whose head is a high school graduate and whose age is between 35 and 55 as of 1983 (1984 survey). For each quintile group of wealth distribution, this table displays the wealth share, ratio of group average to economy-wide average, earnings share, and participation rate.

As can be seen in table 4, the model captures well earnings and wealth differences across quintiles. This is in spite of the fact that it was not estimated to match any of these features nor using data from the PSID. In the data the richest 20 percent of families own nearly 58 percent of total wealth, while in the model they own nearly 59 percent of all wealth. Comparing the model to Chang and Kim (2007), the table shows that this model performs better in capturing the shares of wealth across all quintiles. Most notably, for the second through fourth quantiles, the current model reduces the discrepancy in shares of wealth between model and data.

Finally, the key success of the model, which is absent in Chang and Kim (2007), is the predicted correlation between wealth and participation. Now because of the positive correlation between labor supply and skills the wealthiest display fairly high participation rates. In the data, the fourth and fifth quintiles have labor market participation rates of 87 and 79 percent, respectively. In the baseline model, the fourth quintile participates at a rate of 70 percent while the fifth quintile participates at a rate of 77 percent. In Chang and Kim (2007), these quintiles participate at rates of 50 and 43 percent, respectively. Thus, this model does a considerably better job in capturing the labor supply decision of

the wealthiest.

To summarize, the baseline model replicates well the distributions of average employment rates and average wages as observed in the data. Moreover, it is also consistent with the negative duration dependence of both the hazard rate from employment to non-employment and vice-versa. A final check of the model's consistency shows that its wealth distribution is consistent with salient features of the wealth distribution derived from the PSID. Most importantly, because of the positive correlation between labor supply and skills it generates a realistic correlation between wealth and participation.

Table 4: Summary Statistics of the Wealth Distribution

	Quintile				
	1	2	3	4	5
<i>PSID-primary households</i>					
Share of wealth	1.03	7.07	13.01	21.10	57.76
Group average/population average	0.05	0.36	0.64	1.06	2.97
Share of earnings	14.29	14.67	20.08	25.07	25.86
Participation rate	0.86	0.84	0.83	0.87	0.79
<i>Benchmark Model</i>					
Share of wealth	-2.76	4.36	13.07	25.99	59.34
Group average/population average	-0.14	0.22	0.66	1.29	2.98
Share of earnings	14.89	17.25	18.37	21.21	28.28
Participation rate	0.87	0.75	0.71	0.70	0.77
<i>Chang and Kim (2007)</i>					
Share of wealth	-2.46	3.27	12.21	26.05	60.93
Group average/population average	-0.12	0.16	0.61	1.30	3.08
Share of earnings	13.52	17.87	20.50	22.65	25.46
Participation rate	0.86	0.63	0.56	0.50	0.43

Notes: The PSID statistics reflect the family wealth and earnings in the 1984 survey as reported in Chang and Kim (2007).

6 Implications for the Frisch Elasticity at the Extensive Margin of Labor Supply

This section discusses the model's implications for the Frisch elasticity at the extensive margin of labor supply. First, it presents the baseline model's implied Frisch elasticity. Second, it presents a simple decomposition of this Frisch elasticity by considering two extreme cases of the baseline model.

6.1 Results for the Baseline Model

When the labor supply choice is indivisible the aggregate labor supply elasticity depends on the shape of the reservation wage distribution. Using this distribution the responsiveness of employment can be inferred by looking at the number of individuals with reservation wages near the steady state wage. In the present context, the inverse cumulative distribution of reservation wages for each worker type is constructed using the model's invariant distribution, $\mu_{i,j}(a, x)$, and reservation wages, $x_{i,j}^*(a)$. Next, elasticities for each type are measured by calculating the derivative of this distribution with respect to reservation wages and evaluating it at the type's participation rate. The aggregate elasticity is then calculated as the weighted sum of these elasticities, where the weights equal the employment shares of each type.¹⁷

The implied Frisch elasticity at the extensive margin of labor supply of the model is 0.71. Note that this elasticity reflects no wealth effect as the entire wealth distribution is held constant. For comparison, Chang and Kim (2007) obtain an implied aggregate elasticity of 1.5, while Gourio and Noual (2009) estimate an elasticity of 1.3. Meanwhile, Erosa et al. (2010) obtain an aggregate elasticity (encompassing both intensive and extensive margins) of 1.27. They ar-

¹⁷This procedure is a simple extension of the procedure used by Chang and Kim (2007) to calculate the aggregate labor supply elasticity in the presence of ex-ante heterogeneity across workers.

gue that the extensive margin accounts for 54 percent of this elasticity. Rogerson and Wallenius (2009) find elasticities ranging from 2.25 to 3.0. However, these elasticities also reflect both intensive and extensive margins. While the value of 0.71 is below previous estimates of the Frisch elasticity at the extensive margin, it is still above all estimates of the Frisch elasticity at the intensive margin, which are bounded by 0.60.¹⁸ The fact that the extensive margin responds more to wage changes than the intensive margin is consistent with the observation that over the business cycle changes in aggregate hours are driven more by changes in the number of individuals employed rather than changes in the amount of hours worked per employed individual.¹⁹

Table 5 presents the individual level employment elasticities by worker type. Again, these individual level elasticities reflect the percent change in participation (evaluated at the steady state participation rate for each worker type) given a one percent change in their steady state reservation wage. The results from table 5 show that the individual labor supply elasticity ranges from zero to above 3. In the model, as in the data, a vast majority of the population is employed frequently and hence does not adjust their employment decision. Meanwhile, another portion of the population is employed less frequently and can adjust their labor supply more readily. However, because their contribution to overall employment is small, their elastic response is weighted less. Worth noting is that type d_2 individuals in my model have an elasticity very close to individuals in the models of Chang and Kim (2007) and Gourio and Noyal (2009). However, by disregarding persistent differences in labor supply, their models do not capture separately the very inelastic response of d_1 workers and very elastic response of d_3 workers.

¹⁸See for example, Chetty (2010); Chetty et al. (2009, 2011) or Faberman (2010).

¹⁹See Coleman (1984) and Heckman (1984).

Table 5: Implied Elasticity from the Steady-state Reservation-Wage Distribution, by Worker Type and Aggregate

Disutility of labor	Skills		
	s_3	s_2	s_1
d_1	0.03	0.03	0.03
d_2	1.24	1.17	1.08
d_3	3.86	3.69	3.55
Aggregate	0.71		

Notes: The numbers reflect the elasticity of the labor-market participation rate of each type (and overall) with respect to the reservation wage (evaluated at the steady-state) based on the steady-state reservation wage distribution.

6.2 The Role of Labor Supply Heterogeneity

This subsection presents results for two extreme cases of the baseline model. This is done to understand whether labor supply or skill differences are the main reason for the low implied labor supply elasticity. In the first version of the model, agents only display ex-ante labor supply differences and have equal ex-ante market skills. In the second version of the model, agents only display ex-ante skill differences (akin to Erosa et al., 2010). For each case, the model is estimated using the same data and procedure as the baseline model and imposing the corresponding restriction on skills or labor disutility. In both cases, the models are estimated to match an employment rate of 74.9 percent, the same data target used for the baseline model.

Table 6 presents the implied aggregate labor supply elasticities for each of the three models: labor disutility and skills (baseline); labor disutility only; and skills only. Each elasticity reflects a percentage change in the aggregate labor force participation rate (evaluated at each model's respective steady state employment rate), given a percentage change in the steady state reservation wage holding the entire wealth distribution constant.

The first row reproduces the baseline aggregate elasticity of 0.71. What can be seen from the next two rows is that this low aggregate elasticity is overwhelmingly due to ex-ante labor supply differences. The model where skills are held constant produces an aggregate elasticity of 0.69. The model where labor disutility is held constant produces an aggregate elasticity of 1.12. The key reason behind this result is that the model where labor disutility is held constant does a poor job in replicating the observed differences in average employment rates across workers. Figure 6 presents the distributions of average employment rates from the data (top), model with labor disutility differences (middle), and model with skill differences (bottom).

As can be seen from figure 6, the model with only skill differences produces a distribution of average employment rates which is dense near the steady-state employment rate. Because in the model these employment rate differences translate into reservation wages differences, the reservation wage distribution of this model is dense near a neighborhood of the steady-state wage rate. Hence, a large aggregate labor supply elasticity is recovered. Finally, table 7 shows another dimension where this model fails. Table 7 presents detailed statistics on wealth and earnings for each of the models and the PSID. The model with only skill differences under-performs, relative to the baseline model and model with labor supply differences, in reproducing a realistic wealth distribution and a realistic wealth effect on labor market participation. The model with only skill differences across individuals vastly over-predicts the share of wealth held by the richest 20 percent and under-predicts the share of wealth held by the poorest 20 percent. This further suggests that a model with ex-ante labor supply differences provides a closer description of actual data.

Conversely, the model with only labor supply differences is able to replicate a distribution of average employment rates similar to the one observed in the data. As a consequence, it produces a disperse reservation wage distribution and hence

a low labor supply elasticity. Table 7 shows where the model with only labor disutility differences fails. As can be seen from the table, this model does not have the same wealth effect on labor market participation as the model with both labor disutility and skill differences. In the model with only disutility differences, the richest 20 percent of the population work too little, while the poorest 20 percent work too much. This follows from the fact that in this model, the correlation between average employment and wages is negative. Individuals with a high disutility of labor work only when they receive high enough idiosyncratic productivity shocks and hence their average wage, conditional on employment, is counterfactually high. Because these individuals will have high asset holdings to finance their long non-employment spells, the counterfactual wealth effect on labor market participation is obtained.

Table 6: Aggregate Labor Supply Elasticity by Model

Model	Elasticity
Benchmark Model	0.71
Labor disutility only	0.69
Skills only	1.12

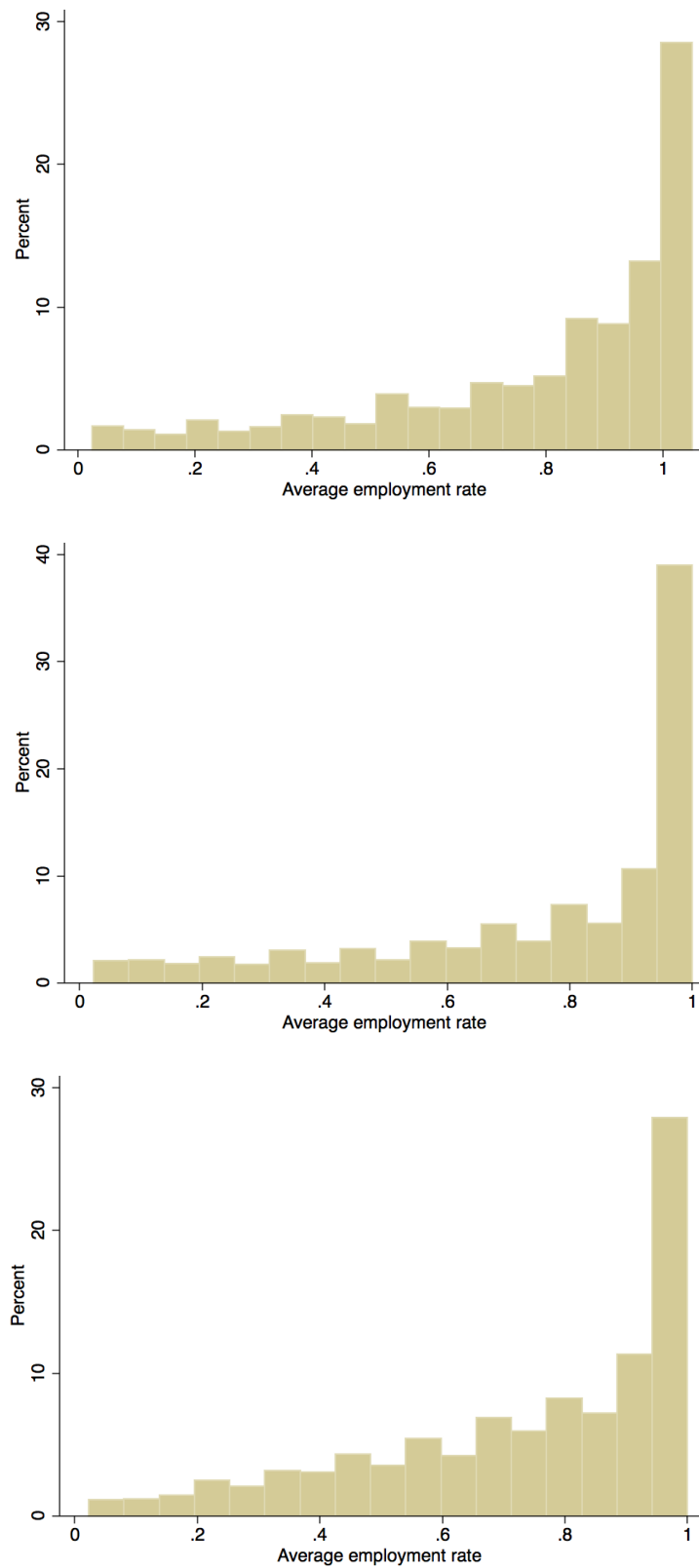


Figure 6: Distributions of Average Employment Rates: Data (top), Model with Labor Disutility (middle), and Model with Skills (bottom).

Table 7: Summary Statistics of the Wealth Distribution

	Quintile				
	1	2	3	4	5
<i>PSID-primary households</i>					
Share of wealth	1.03	7.07	13.01	21.10	57.76
Group average/population average	0.05	0.36	0.64	1.06	2.97
Share of earnings	14.29	14.67	20.08	25.07	25.86
Participation rate	0.86	0.84	0.83	0.87	0.79
<i>Benchmark Model</i>					
Share of wealth	-2.76	4.36	13.07	25.99	59.34
Group average/population average	-0.14	0.22	0.66	1.29	2.98
Share of earnings	14.89	17.25	18.37	21.21	28.28
Participation rate	0.87	0.75	0.71	0.70	0.77
<i>Labor disutility only</i>					
Share of wealth	-1.74	4.31	13.08	26.25	58.11
Group average/population average	-0.09	0.22	0.65	1.31	2.91
Share of earnings	14.84	19.01	20.93	22.93	23.16
Participation rate	0.92	0.83	0.78	0.72	0.62
<i>Skills only</i>					
Share of wealth	-5.78	0.54	10.37	25.21	69.66
Group average/population average	-0.29	0.03	0.52	1.26	3.48
Share of earnings	14.86	18.02	19.35	21.05	26.72
Participation rate	0.96	0.81	0.73	0.66	0.57

Notes: The PSID statistics reflect the family wealth and earnings in the 1984 survey as reported in Chang and Kim (2007).

7 Robustness

This section presents alternative specifications of the model to verify the robustness of the baseline Frisch elasticity at the extensive margin of 0.71. In particular, these exercises are intended to quantify how important is the assumption of only allowing for three skill and three labor disutility types. To do so, the baseline model is re-specified to allow for five skill types and five labor disutility types; i.e. a 5×5 rather than 3×3 model. This augmented model is estimated again using the procedure described in section 4 and the model's implied Frisch elasticity is calculated. The results of this exercise appear in table 8.

The implied elasticity from the 5×5 model is 0.62, nearly 13% smaller than the elasticity from the baseline 3×3 model of 0.71. The reason for the drop in the extensive margin elasticity can be seen in table 9 which presents the estimation results for this model. In the 5×5 model, individuals with high skills and low labor disutility (the top left corner of the matrix of population proportions) contribute disproportionately to the overall population. Because of their high labor supply they also contribute disproportionately to aggregate employment in a very inelastic fashion.

To quantify how much of this drop in the recovered elasticity is due to the increase in skill versus labor disutility heterogeneity, an alternative version of the augmented model where only labor disutility differences are present is estimated. The implied elasticity of this 1×5 model is 0.52 and thus explains all of the decline in the Frisch elasticity when moving from a 3×3 to 5×5 model. Compared to the 5×5 model, the 1×5 model recovers and even smaller Frisch elasticity because it violates the positive correlation between average wages and employment observed in the data (-0.24 versus 0.39).²⁰

To conclude, the results from this section suggest that the Frisch elasticity of the extensive margin is likely lower than the baseline value of 0.71. In other

²⁰Note that is also true in the version of the model with only three skill types.

words, the assumption of only allowing for three skill and three disutility types in the model is not without loss of generality. Furthermore, when comparing these results to those obtained in section 6.2, they suggest that as the degree of labor disutility heterogeneity increases, the model's implied Frisch elasticity at the extensive margin decreases.

Table 8: Aggregate Labor Supply Elasticity by Model, Alternative Specifications

Model	Elasticity
Benchmark model (3 skills x 3 disutilities)	0.71
Augmented benchmark model (5 skills x 5 disutilities)	0.62
Augmented labor disutility only (5 disutilities)	0.52

Table 9: Estimated Parameter Values Augmented Model

	Skills				
Disutility of labor	$s_5 = 1.00^\dagger$	$s_4 = 0.56$	$s_3 = 0.41$	$s_2 = 0.30$	$s_1 = 0.25$
$d_1 = 0.21$	$p_{15} = 0.44$	$p_{14} = 0.01$	$p_{13} = 0.01$	$p_{12} = 0.01$	$p_{11} = 0.00$
$d_2 = 0.73$	$p_{25} = 0.05$	$p_{24} = 0.05$	$p_{23} = 0.03$	$p_{22} = 0.00$	$p_{21} = 0.00$
$d_3 = 0.93$	$p_{35} = 0.01$	$p_{34} = 0.03$	$p_{33} = 0.18$	$p_{32} = 0.01$	$p_{31} = 0.00$
$d_4 = 1.56$	$p_{45} = 0.00$	$p_{44} = 0.00$	$p_{43} = 0.04$	$p_{42} = 0.03$	$p_{41} = 0.02$
$d_5 = 3.99$	$p_{55} = 0.00$	$p_{54} = 0.00$	$p_{53} = 0.00$	$p_{52} = 0.00$	$p_{51} = 0.05$
	Value				
ρ_x	0.92				
σ_ϵ	0.20				

Notes: \dagger by normalization. Discount factor $\beta = 0.98742$, found from capital market clearing.

8 Conclusion

This paper examines the role of ex-ante heterogeneity across workers in determining the Frisch elasticity at the extensive margin of employment. Motivated by empirical observations from the NLSY that show large differences in average employment rates across individuals that do not project on wages, a heterogeneous agent model with incomplete markets and indivisible labor supply is presented to match these facts. The novel ingredients of the model are allowing agents to differ in their disutility of labor and market skills, both of which remain fixed across time. Unlike most of the previous literature, with Erosa et al. (2010) as an important exception, the model allows for a rich description of ex-ante heterogeneity (labor disutility and skills), and ex-post heterogeneity (idiosyncratic productivity shocks and assets) across agents. Rather than calibrating the model to match aggregate moments, the model is estimated using indirect inference with key micro-level parameters.

The main result of the paper is summarized as follows. Once agents display a realistic amount of ex-ante heterogeneity in labor supply and skills, a very large macro-level elasticity is no longer obtained through the extensive margin of labor supply. The implied aggregate labor supply elasticity of the model is 0.71. Robustness exercises generate elasticities as low as 0.62 and suggest that as the degree of heterogeneity in the model increases, the elasticity decreases. These elasticities are below previous extensive margin estimates (typically above 1) and above estimates of the elasticity at the intensive margin (at most 0.60), which contributes less (relative to the extensive margin) to changes in aggregate employment over the business cycle relative.

A simple decomposition reveals the importance of these labor supply differences for the inferred Frisch elasticity of the extensive margin. In a version of the model with no ex-ante labor supply differences (akin to Erosa et al., 2010), the recovered elasticity is 1.1, nearly 53 percent larger than the elasticity obtained

from the baseline model. Meanwhile, in a version of the model with no ex-ante skill differences the recovered elasticity is 0.69, which is virtually identical to the elasticity obtained from the baseline model. This version, however, violates the positive cross-sectional correlation between average employment rates and average wages that is observed in the data. In the baseline model with labor supply and skill differences, this correlation is positive. Because of this correlation it generates a realistic wealth effect on labor market participation, which is not found in the literature.

Future research should consider allowing for some intensive margin adjustment (e.g., choice of hours conditional on being employed subject to some minimum requirement) as an extension of the present setting to verify that the results are not driven by the assumption of no intensive margin choice. Verifying that the hours choice by worker type is consistent with what is observed in the data is another important check of my model's consistency. Allowing for a distinction between men and women in the model is also a promising venue of research as the current model abstracts from the difference between the labor market participation decision of a married woman versus a single man. Work by Guner et al. (2008) shows that this distinction is very important. Finally, extending the model to allow for business cycle shocks is also a promising direction of research. The structure of the model can help quantify how much of the volatility of aggregate employment and wages is due to the employment response of each of the worker types over the business cycle. Obtaining answers to these questions will further our knowledge about the aggregate implications of individual level heterogeneity.

References

- Abowd, John M. and David Card**, “On the Covariance Structure of Earnings and Hours Change,” *Econometrica*, 1989, 57 (2), 411–45.
- Aiyagari, S. Rao**, “Uninsured Idiosyncratic Risk and Aggregate Saving,” *The Quarterly Journal of Economics*, 1994, 109 (3), 659–84.
- Altonji, Joseph, Anthony A. Smith, and Ivan Vidangos**, “Modeling Earnings Dynamics,” 2009. NBER Working Paper # 14743.
- Altonji, Joseph G.**, “Intertemporal Substitution in Labor Supply: Evidence from Micro Data,” *Journal of Political Economy*, 1986, 94 (3, part 2), S176–S215.
- Auerbach, Alan J. and Laurence J. Kotlikoff**, *Dynamic Fiscal Policy*, Cambridge, UK: Cambridge University Press, 1987.
- Bils, Mark, Yongsung Chang, and Sun-Bin Kim**, “Comparative Advantage and Unemployment,” 2009. Mimeo University of Rochester.
- Chang, Yongsung and Sun-Bin Kim**, “From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy,” *International Economic Review*, 2006, 47 (1), 1–27.

— **and** —, “Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations,” *American Economic Review*, 2007, 97 (5), 1939–56.

Chetty, Raj, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” 2010. NBER Working Paper # 15616.

—, **John N. Friedman, Tore Olsen, and Luigi Pistaferri**, “Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records,” 2009. NBER Working Paper # 15617.

—, —, —, **and** —, “Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records,” 2011. Forthcoming *Quarterly Journal of Economics*.

Cho, Jang-Ok, “Ex post heterogeneity and the business cycle,” *Journal of Economic Dynamics and Control*, 1995, 19, 533–51.

Coleman, Thomas, “Essays on Aggregate Labor Market Business Cycle Fluctuations.” PhD dissertation, University of Chicago 1984.

Erosa, Andrés, Luisa Fuster, and Gueorgui Kambourov, “A Micro-founded Theory of Aggregate Labor Supply with Non-linear Wages,” 2010. Mimeo University of Toronto.

Faberman, E. Jason, “Revisiting the Role of Home Production in Life-Cycle Labor Supply,” 2010. Research Department, Federal Reserve Bank of Philadelphia, Working Paper 10-3.

French, Eric B., “The Effects of Health, Wealth and Wages on Labour Supply and Retirement Behaviour,” *Review of Economic Studies*, 2005, 72 (2), 395–427.

- Gallant, A. Ronald and George Tauchen**, “Indirect Inference,” *Journal of Applied Econometrics*, 1993, 8, S85–S118.
- Ghez, Gilbert and Gary Becker**, *The Allocation of Time and Goods over the Life Cycle*, New York: National Bureau of Economic Research, Columbia University Press, 1975.
- Gourieroux, C, A Monfort, and E Renault**, “Indirect Inference,” *Journal of Applied Econometrics*, 1993, 8, S85–S118.
- Gourio, François and Pierre-Alexandre Noul**, “The Marginal Worker and The Aggregate Elasticity of Labor Supply,” 2009. Boston University Working Paper WP2006-009.
- Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura**, “Taxation, Aggregates and the Household,” 2008. IZA Discussion Paper No. 3318.
- Guvenen, Fatih and Anthony A. Smith**, “Inferring Labor Income Risk From Economic Choices: An Indirect Inference Approach,” 2010. Working paper.
- Hansen, Gary D.**, “Indivisible Labor and The Business Cycle,” *Journal of Monetary Economics*, 1985, 16, 309–27.
- Heckman, James**, “Comments on Ashenfelter and Kydland Papers,” *Carnegie-Rochester Conference Series on Public Policy*, 1984, 21, 209–24.
- Keane, Michael and Anthony A. Smith**, “Generalized Indirect Inference for Discrete Choice Models,” 2003. Working paper.
- Keane, Michael P. and Kenneth I. Wolpin**, “Eliminating Race Differences in School Attainment and Labor Market Success,” *Journal of Labor Economics*, 2000, 18 (4), 614–652.
- **and Richard Rogerson**, “Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective,” 2011. NBER Working Paper # 17430.

- Kimmel, Jean and Thomas J. Kniesner**, “New evidence on labor supply: Employment versus hours elasticities by sex and marital status,” *Journal of Monetary Economics*, 1998, *42*, 289–01.
- Krusell, Per, Toshihiko Mukoyama, Richard Rogerson, and Ayşegül Şahin**, “A Three State Model of Worker Flows in General Equilibrium,” *Journal of Economic Theory*, 2011, *146* (3), 1107–1133.
- Kydland, Finn E. and Edward C. Prescott**, “Time to Build and Aggregate Fluctuations,” *Econometrica*, 1982.
- Lucas, Robert E. Jr. and Leonard Rapping**, “Real Wages, Employment and Inflation,” *Journal of Political Economy*, 1969.
- MaCurdy, Thomas E.**, “An Empirical Model of Labor Supply in a Life-Cycle Setting,” *Journal of Political Economy*, 1981.
- Prescott, Edward C.**, “Why Do Americans Work So Much More Than Europeans?,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 2004, *28* (1), 2–13.
- Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery**, *Numerical Recipes in C*, Cambridge: Cambridge University Press, 1992.
- Ríos-Rull, José-Víctor**, “Computation of Equilibria in Heterogeneous-Agent Models,” in Ramon Marimon and Andrew Scott, eds., *Computational Methods for the Study of Dynamic Economies*, New York: Oxford University Press, 1999.
- Rogerson, Richard**, “Indivisible Labor, Lotteries and Equilibrium,” *Journal of Monetary Economics*, 1988, *21*, 3–16.

- **and Johanna Wallenius**, “Micro and Macro Elasticities in a Life Cycle Model with Taxes,” *Journal of Economic Theory*, 2009, *144* (6), 2277–92.
- Smith, Anthony A.**, “Estimating Nonlinear Time-Series Models using Simulated Vector Autoregressions,” *Journal of Applied Econometrics*, 1993, *8*, S63–S84.
- , “Indirect Inference,” in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, 2nd ed., Palgrave Macmillan, 2008. The New Palgrave Dictionary of Economics Online. Palgrave Macmillan. 22 September 2010 <http://www.dictionaryofeconomics.com/article?id=pde2008_I000259> doi:10.1057/9780230226203.0778.
- Smith, Anthony A. Jr.**, “Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models.” PhD dissertation, Duke University 1990.
- Tauchen, George.**, “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions,” *Economics Letters*, 1986.

Data Appendix

A Data

A.1 Linking Employers Across Survey Years

The NLSY allows the linking of an individual’s job reports across consecutive survey years. In linking reports across survey years the method suggested in the NLSY technical Appendix # is followed by using the variables defined as “*Previous job number at last interview #1-5*”.

A.2 Constructing Quarterly Employment Status

The NLSY79 provides variables containing the weekly employment status of each individual in the sample in their work history file. These variables are named “*Labor Force Status Week #*”, where $\#$ serves as a place holder for the week number in question. Each calendar week is assigned a number starting with 1 (corresponding to the first of January 1978), through 1531 (corresponding to the week starting with February 29th 2007). Quarterly employment status for each individual is constructed as follows:

1. For quarter q determine the week numbers \underline{w} and \overline{w} which correspond to the first and last weeks in the quarter.
2. For each week in $[\underline{w}, \overline{w}]$ check if the individual is employed (status code ≥ 100 or 3), non-employed (status code 2, 4, or 5) or missing (status code 0 or 7).
3. If the individual is employed for at least 7 weeks in the quarter, set her quarterly employment status to employed. If the individual is not employed for at least 7 weeks, but has at least one week where her status is not missing, set her quarterly employment status to non-employed. Otherwise, set her status to missing.

A.3 Wages

Hourly wage rates are taken from the variables “*Hours usually worked at current/most recent job*” and “*Hourly Rate of Pay Job #1-5*”. From 1979-1993 detailed information on the CPS or current/most recent employer is collected in the CPS section, while after 1993 the CPS employer is always the first job coded. Hence, for survey years 1979-1993 it is necessary to look at both sets of variables to obtain complete information on the CPS job. If an individual reports wages in

units other than hourly, the NLSY calculates an hourly wage rate based on the earnings reported, the unit in which they are reported and usual hours worked on the job. Nominal wages are deflated using the Consumer Price Index for all urban consumers and all items (CPI-U), which is seasonally adjusted. Missing wages are imputed using the previous or next wage report from the same job, if available.

A.4 Hours

To identify hours worked in each job the variables “*Hours usually worked at current/most recent job*” and “*Hours per week usually worked at Job # 1-5*” are combined, deferring to the CPS report whenever the job coincides with the current/most recent employer.

A.5 Definition of Quarterly Wage

The quarterly wage rate is defined as the hourly wage rate of the job the individual works at the most during the quarter in question. Time spent working at the job is measured as the product of hours per week times weeks worked in the quarter.

B Computation of the Steady-State Equilibrium

The computational strategy used to compute the steady-state equilibrium of the model is an extension of the one used in Chang and Kim (2007) to take into account multiple worker types. As in Ríos-Rull (1999), the goal is to find the discount rate β that clears the capital market given an interest rate of 1%. The algorithm proceeds as follows:

0. Initialize guesses (or current estimates) for $\{s_1, \dots, s_{N_s}\}, \{d_1, \dots, d_{N_d}\}, \{p_{sd}\}_{s=1, d=1}^{N_s, N_d}, \sigma_x, \rho_x$.

1. Choose the grid points for asset holdings a and idiosyncratic productivity x . Denote the number of grids by N_a and N_x . I set $N_a = 1,666$ and $N_x = 10$. Asset holdings a are restricted to the range $[-2, 2000]$, where the average asset holdings are 13.7. The grid points on asset are not equally spaced; more points are assigned on the bottom of the asset range to better approximate the savings decisions of workers with lower assets. For idiosyncratic productivity, construct a vector of length N_x , whose elements $\ln x_j$, are equally spaced on the interval $[-3\sigma_x/\sqrt{1-\rho_x^2}, +3\sigma_x/\sqrt{1-\rho_x^2}]$. I use Tauchen's (1986) algorithm to approximate the idiosyncratic productivity process using a transition matrix.

2. Given values for β , skills s , and labor disutilities d , solve for the value functions $\{V_{sd}^E, V_{sd}^N, V_{sd}\}_{s=1, d=1}^{N_s, N_d}$ at each grid point of the individual states. This also yields the optimal decision rules for asset holdings and labor supply for each worker type $\{a'_{sd}(a_i, x_j), h_{sd}(a_i, x_j)\}_{s=1, d=1}^{N_s, N_d}$. The value functions are found iteratively as follows:

(a) Initialize the value functions $V_{sd}^E(a_i, x_j)$ and $V_{sd}^N(a_i, x_j)$ for all $i = 1, \dots, N_a, j = 1, \dots, N_x, s = 1, \dots, N_s$ and $d = 1, \dots, N_d$.

(b) Obtain updated guesses of the value functions by evaluating the dis-

cretized versions

$$\begin{aligned}\tilde{V}_{sd}^E(a_i, x_j) &= \max_{a' \in \{a_1, \dots, a_{N_a}\}} \left\{ \ln \left(w\bar{h}s_s x_j + (1+r)a_i - a' \right) - d_d \right. \\ &\quad \left. + \beta \sum_{k=1}^{N_x} V_{sd}(a', x_j) \pi_x(x_k | x_j) \right\} \\ \tilde{V}_{sd}^N(a_i, x_j) &= \max_{a' \in \{a_1, \dots, a_{N_a}\}} \left\{ \ln \left((1+r)a_i - a' \right) \right. \\ &\quad \left. + \beta \sum_{k=1}^{N_x} V_{sd}(a', x_j) \pi_x(x_k | x_j) \right\}\end{aligned}$$

where $\pi_x(x' | x_j)$ is the transition probability of x_j to x' . Update $\tilde{V}_{sd}(a_i, x_j) = \max\{\tilde{V}_{sd}^E(a_i, x_j), \tilde{V}_{sd}^N(a_i, x_j)\}$.

- (c) If \tilde{V} and V are close enough for all grid points and for each s, d pair, then we have found the value functions. Otherwise, set $V_{sd}^E = \tilde{V}_{sd}^E$ for each s, d pair and all grid points (and similarly for V^N), and go back to step 2 (b).

3. Using $\{a'_{sd}(a_i, x_j)\}_{s=1, d=1}^{N_s, N_d}$ obtained from step 2 and $\pi_x(x' | x_j)$, obtain the time-invariant measures $\{\mu_{sd}^*(a_i, x_j)\}_{s=1, d=1}^{N_s, N_d}$ as follows:

- (a) Initialize the measures $\{\mu_{sd}(a_i, x_j)\}_{s=1, d=1}^{N_s, N_d}$, such that

$\sum_{i=1}^{N_a} \sum_{j=1}^{N_x} \mu_{sd}(a_i, x_j) = p_{sd}$, where p_{sd} is the proportion of the population with skills $s = s_s$ and labor disutility $d = d_d$.

- (b) Update each measure by evaluating a discretized version of (5) (for each (s, d) pair):

$$\mu'_{sd}(a_{i'}, x_{j'}) = \sum_{i=1}^{N_a} \sum_{j=1}^{N_x} \mathbf{1}_{a_{i'} = a'_{sd}(a_i, x_j)} \mu_{sd}(a_i, x_j) \pi_x(x_{j'} | x_j)$$

- (c) If μ'_{sd} and μ_{sd} are close enough for all grid points and each s, d pair, then we have found the time-invariant measure. Otherwise, set $\mu_{sd} = \mu'_{sd}$ and go back to step 3 (b).

4. Calculate the real interest rate as a function of β , $r(\beta) = \alpha \left(K(\beta), L(\beta) \right)^{1-\alpha} - \delta$, where

$$K(\beta) = \sum_{s=1}^{N_s} \sum_{d=1}^{N_d} \sum_{i=1}^{N_a} \sum_{j=1}^{N_x} a_i \mu_{sd}^*(a_i, x_j)$$

and

$$L(\beta) = \sum_{s=1}^{N_s} \sum_{d=1}^{N_d} \sum_{i=1}^{N_a} \sum_{j=1}^{N_x} s_s x_j h_{sd}(a_i, x_j) \mu_{sd}^*(a_i, x_j).$$

If $r(\beta)$ is close enough to the assumed value of the real interest rate, we have found the steady-state. Otherwise, choose another β and go back to step 2.

C Estimates from the Auxiliary Model

Table 10: Estimated Results for Auxiliary Model: Actual vs Model-Generated Data

Equation										
Parameter	$E_{it}E_{it-1}$		$E_{it}(1 - E_{it-1})$		$(1 - E_{it})E_{it-1}$		$(1 - E_{it})(1 - E_{it-1})$		W	
	Actual	Model	Actual	Model	Actual	Model	Actual	Model	Actual	Model
γ_{ED}	0.14111985	0.12240	-.05006039	-0.04476	.00434902	-0.00356	-.09540847	-0.07407	.01474449	0.01862
γ_{ND}	-0.25734125	-0.23632	.02473657	0.01476	-.04687025	-0.06382	.27947493	0.28538	-.00817502	-0.03134
γ_w	-0.25734125	-0.03348	.01587276	0.04549	-.03097592	-0.06282	.06152133	0.05081	.80257107	0.78664
$\gamma_{\bar{e}}$	0.00640665	0.13790	.17307713	0.12332	-.1654487	-0.20591	-.01403509	-0.05531	-.10842342	-0.15699
$\gamma_{\bar{w}}$.04605656	0.04321	-.02025478	-0.05656	.01423726	0.03484	-.04003905	-0.02148	.14583943	0.17233
γ_0	0.58185589	0.49731	-.02004111	0.01560	.17160448	0.22762	.26658075	0.25947	.07126272	0.10718
R^2	0.7490	0.75829	0.0959	0.09548	0.0507	0.09299	0.7633	0.74	0.8338	0.8334

Notes: All coefficients are significant at the 1% level. Data coefficients are weighted averages across regression coefficients by age groups [25, 30), [30, 35), [35, 40), [40, 48). Data standard errors are also weighted averages of standard errors by age group regression. Model coefficients are averages over 100 simulations. Model standard errors are calculated from the distribution of each parameter over the 100 simulations.

Table 11: Estimated Results for Other Moments: Actual vs Model-Generated Data

Moment	Actual Data	Model Simulated Data
μ_e	0.746	0.757 (0.0008)
μ_w	-0.066	-0.058 (0.001)
σ_e	0.281	0.291 (0.0003)
σ_w	0.520	0.450 (0.0007)
$Skewness_e$	-0.989	-0.987 (0.005)
$Skewness_w$	0.145	-0.07 (0.004)
$kurt_e$	2.766	2.70 (0.012)
$kurt_w$	3.155	3.038 (0.008)
$\rho(e, w)$	0.391	0.323 (0.002)

Notes: Model standard errors in parentheses. Model moments are averages over 100 simulations. Model standard errors are calculated from the distribution of each moment over the 100 simulations.